Linear Algebra and its Applications 439 (2013) 2584–2597

Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Diagonals and numerical ranges of direct sums of matrices

Hsin-Yi Lee [1]

Department of Applied Mathematics, National Chiao Tung University, Hsinchu 30010, Taiwan, ROC

ARTICLE INFO

Article history:
Received 4 March 2013
Accepted 13 July 2013
Available online 1 August 2013
Submitted by L. Rodman

MSC:
15A60

Keywords:
Numerical range
Direct sum
Compression

ABSTRACT

For any n-by-n matrix A, we consider the maximum number $k = k(A)$ for which there is a k-by-k compression of A with all its diagonal entries in the boundary $\partial W(A)$ of the numerical range $W(A)$ of A. If A is a normal or a quadratic matrix, then the exact value of $k(A)$ can be computed. For a matrix A of the form $B \oplus C$, we show that $k(A) = 2$ if and only if the numerical range of one summand, say, B is contained in the interior of the numerical range of the other summand C and $k(C) = 2$. For an irreducible matrix A, we can determine exactly when the value of $k(A)$ equals the size of A. These are then applied to determine $k(A)$ for a reducible matrix A of size 4 in terms of the shape of $W(A)$.
© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Let A be an n-by-n complex matrix. Its numerical range $W(A)$ is, by definition, the set $\{\langle Ax, x\rangle \colon x \in \mathbb{C}^n, \ \|x\| = 1\}$, where $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ denote the standard inner product and its associated norm in $\mathbb{C}^n$, respectively. It is well known that $W(A)$ is a nonempty compact convex subset of the complex plane. For other properties of the numerical range, we refer the reader to [4, Chapter 1]. Let $k(A)$ be the maximum number k of orthonormal vectors $x_1, \ldots, x_n \in \mathbb{C}^n$ with $\langle Ax_j, x_j\rangle$ in the boundary $\partial W(A)$ of $W(A)$ for all j. Note that $k(A)$ is also the maximum size of a compression of A with all its diagonal entries in $\partial W(A)$. Recall that a k-by-k matrix B is a *compression* of A if $B = V^*AV$ for some n-by-k matrix V with $V^*V = I_k$. In particular, if n equals k, then A and B are said to be *unitarily similar*, which we denote by $A \cong B$. The number $k(A)$ was introduced in [3] and [7]. It relates properties of the numerical range to the compressions of A. In particular, it was shown in

E-mail address: hylee.am95g@nctu.edu.tw.

[1] The contents of this paper form a part of the author's Ph.D. dissertation at National Chiao Tung University under the supervision of Pei Yuan Wu. The author is grateful to the latter for the guidance and support provided over the years to have this work accomplished.

0024-3795/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.laa.2013.07.019

[3, Lemma 4.1 and Theorem 4.4] that $2 \leqslant k(A) \leqslant n$ for any n-by-n $(n \geqslant 2)$ matrix A, and $k(A) = \lceil n/2 \rceil$ for any S_n matrix A $(n \geqslant 3)$. Recall that an n-by-n matrix A is of class S_n if it is a *contraction*, that is, $\|A\| \equiv \max_{\|x\|=1} \|Ax\| \leqslant 1$, its eigenvalues are all in the open unit disc $\mathbb{D} \equiv \{z \in \mathbb{C}\colon |z| \leqslant 1\}$, and the rank of $I_n - A^*A$ equals one. In [7, Theorem 3.1], it was proven that $k(A) = n$ for an n-by-n $(n \geqslant 2)$ weighted shift matrix A with weights $w_1, \dots, w_n$ if and only if either $|w_1| = \cdots = |w_n|$ or n is even and $|w_1| = |w_3| = \cdots = |w_{n-1}|$ and $|w_2| = |w_4| = \cdots = |w_n|$. Recall that an n-by-n $(n \geqslant 2)$ matrix of the form

$$\begin{bmatrix} 0 & w_1 & & \\ & 0 & \ddots & \\ & & \ddots & w_{n-1} \\ w_n & & & 0 \end{bmatrix}$$

is called a *weighted shift matrix* with *weights* $w_1, \dots, w_n$.

In Section 2 below, we first determine the value of $k(A)$ for a normal matrix A (Proposition 2.1). Then we consider the direct sum $A = B \oplus C$, where the numerical ranges $W(B)$ and $W(C)$ are assumed to be disjoint. In this case, we show that the value of $k(A)$ is equal to the sum of $k_1(B)$ and $k_1(C)$ (Theorem 2.2), where $k_1(B)$ and $k_1(C)$ are defined as follows. We define $k_1(B)$ to be the maximum number k for which there are orthonormal vectors $x_1, \dots, x_k$ in $\mathbb{C}^n$ such that $\langle Bx_i, x_i \rangle$ is in $\partial W(A) \cap \partial W(B)$ for all $i = 1, \dots, k$, and similarly for $k_1(C)$. Based on the proof of Theorem 2.2, we obtain the same formula for $k(A)$ under a slightly weaker condition on B and C (Theorem 2.4). In Section 3, we give some applications of Theorem 2.4. The first one (Proposition 3.1) shows that the equality $k(A) = k_1(B) + k_1(C)$ holds for a matrix A of the form $B \oplus C$ with normal C. In particular, we are able to determine the value of $k(A)$ for any 4-by-4 reducible matrix A (Corollary 3.4 and Propositions 3.7–3.9). Moreover, the number $k(A \oplus (A + aI_n))$ can be determined for any n-by-n matrix A and nonzero complex number a (Proposition 3.10). At the end of Section 3, we propose several open questions on $k(B \oplus C)$ and give a partial answer for one of them (Proposition 3.11). That is, the equality $k(\bigoplus_{j=1}^m A) = m \cdot k(A)$ holds if the dimension of $H_\xi(A)$ equals one for each $\xi \in \partial W(A)$, where the subspace $H_\xi(A)$ is defined in the first paragraph of Section 2. By using this, we can determine the value of $k(A)$ for a quadratic matrix A (Corollary 3.12). Recall that a *quadratic matrix* A is one which satisfies $A^2 + z_1 A + z_2 I = 0$ for some scalars z_1 and z_2.

We end this section by fixing some notation. For any finite square matrix A, we use $\operatorname{Re} A = (A + A^*)/2$ and $\operatorname{Im} A = (A - A^*)/(2i)$ to denote its *real* and *imaginary parts*, respectively. The set of eigenvalues of A is denoted by $\sigma(A)$. A is called *positive definite*, denoted by $A > 0$, if A is Hermitian and $\langle Ax, x \rangle > 0$ for all $x \neq 0$. I_n is the n-by-n identity matrix. The n-by-n diagonal matrix with diagonals $\xi_1, \dots, \xi_n$ is denoted by $\operatorname{diag}(\xi_1, \dots, \xi_n)$. The *cardinal number* of a set S is $\#(S)$. The notation δ_{ij} is the *Kronecker delta*, i.e., δ_{ij} has the value 1 if $i = j$, and the value 0 if otherwise. The *span* of a nonempty subset S of a vector space V, denoted by $\operatorname{span}(S)$, is the subspace consisting of all linear combinations of the vectors in S.

2. Direct sum

We start by reviewing a few basic facts concerning the boundary points of a numerical range. For an n-by-n matrix A, a point ξ in $\partial W(A)$ and a supporting line L of $W(A)$ which passes through ξ, there is a θ in $[0, 2\pi)$ such that the ray from the origin which forms angle θ from the positive x-axis is perpendicular to L. In this case, $\operatorname{Re}(e^{-i\theta}\xi)$ is the maximum eigenvalue of $\operatorname{Re}(e^{-i\theta}A)$ with the corresponding eigenspace $E_{\xi,L}(A) \equiv \ker \operatorname{Re}(e^{-i\theta}(A - \xi I_n))$. Let $K_\xi(A)$ denote the set $\{x \in \mathbb{C}^n\colon \langle Ax, x \rangle = \xi \|x\|^2\}$ and $H_\xi(A)$ the subspace spanned by $K_\xi(A)$. If the matrix A is clear from the context, we will abbreviate these to $E_{\xi,L}$, K_ξ and H_ξ, respectively. For other related properties, we refer the reader to [2, Theorem 1] and [7, Proposition 2.2]. The next proposition on the value of $k(A)$ for a normal matrix A is an easy consequence of [7, Lemma 2.9]. It can be regarded as a motivation for our study of this topic.

Proposition 2.1. *If A is an n-by-n normal matrix with p eigenvalues (counting multiplicity) in $\partial W(A)$, then $k(A) = p$.*

Proof. We may assume, after a unitary similarity, that A is a matrix of the form $B \oplus C$, where $B = \operatorname{diag}(\lambda_1, \ldots, \lambda_p)$ and $C = \operatorname{diag}(\lambda_{p+1}, \ldots, \lambda_n)$ with $\lambda_1, \ldots, \lambda_p \in \partial W(A)$ and $\lambda_{p+1}, \ldots, \lambda_n \in \operatorname{int} W(B)$. It follows from [7, Lemma 2.9] that $k(A) = k(B \oplus C) = k(B) = p$. □

One of our main results of this section is the following theorem for $k(A)$ when A is a matrix of the form $B \oplus C$ with disjoint $W(B)$ and $W(C)$. Recall that the value of $k_1(B)$ is the maximum number k for which there are orthonormal vectors $x_1, \ldots, x_k$ in $\mathbb{C}^n$ such that $\langle Bx_i, x_i\rangle$ is in $\partial W(A) \cap \partial W(B)$ for all $i = 1, \ldots, k$. If the subset $\partial W(A) \cap \partial W(B)$ is empty, then we define $k_1(B) = 0$. The following theorem provides a formula for determining the value of $k(A)$ by $k_1(B)$ and $k_1(C)$.

Theorem 2.2. *Let $A = B \oplus C$, where B and C are n-by-n and m-by-m matrices, respectively. If the numerical ranges $W(B)$ and $W(C)$ are disjoint, then $k(A) = k_1(B) + k_1(C) \leqslant k(B) + k(C)$. In this case, $k(A) = k(B) + k(C)$ if and only if $k_1(B) = k(B)$ and $k_1(C) = k(C)$. In particular, $k(A) = m + n$ if and only if $k_1(B) = k(B) = m$ and $k_1(C) = k(C) = n$.*

This will be proven after the following lemma which is the case when C corresponds to a 1-by-1 matrix $[c]$.

Recall that z is an *extreme point* of the convex subset Δ of $\mathbb{C}$ if z belongs to Δ and is not expressible as a convex combination of two other (distinct) points of Δ; otherwise, z is a *nonextreme point*. Recall also that a point z is called a *corner* of a convex set Δ of the complex plane if z is in the closure of Δ and Δ has two supporting lines passing through z. If $\xi = \langle Ax, x\rangle$ and $\|x\| = 1$, then x is called a unit vector corresponding to ξ.

Lemma 2.3. *If $A = B \oplus [c]$ is an n-by-n matrix, where B is of size $n - 1$ and c is a scalar, then $k(A) = k_1(B) + k_1([c])$.*

Proof. By Proposition 2.1, we may assume that the interior of the numerical range $W(B)$ is nonempty. If c is in the interior of $W(B)$, then $k(A) = k(B)$ by [7, Lemma 2.9]. Obviously, $k(B) = k_1(B)$ and $k_1([c]) = 0$ in this case. Hence it remains to consider the case when c is outside the interior of $W(B)$. That is, we will prove that $k(A) = k_1(B) + 1$ for $c \notin \operatorname{int} W(B)$. By the definition of $k(A)$, there exist $\xi_j = \langle Az_j, z_j\rangle \in \partial W(A)$, $j = 1, 2, \ldots, k(A)$, where $z_j = x_j \oplus y_j$, and $\langle z_i, z_j\rangle = \delta_{ij}$ for $i, j = 1, \ldots, k(A)$. Clearly, the inequality $k(A) \geqslant k_1(B) + 1$ holds. Assume that $k(A) \geqslant k_1(B) + 2$. We claim that every x_j is a nonzero vector. Indeed, if $x_{j_0} = 0$ for some j_0, then $y_{j_0} \neq 0$ and $\langle z_j, z_{j_0}\rangle = \langle y_j, y_{j_0}\rangle = 0$ for all $j \neq j_0$. This implies that $y_j = 0$ for all $j \neq j_0$ and thus $k_1(B)$ is at least $k_1(B) + 1$, which is absurd. Hence the claim has been proven. From $\xi_j = \langle Az_j, z_j\rangle = \|x_j\|^2 b_j + \|y_j\|^2 c \in \partial W(A)$, where $b_j = \langle B(x_j/\|x_j\|), x_j/\|x_j\|\rangle$, it follows that ξ_j is in the (possibly degenerate) line segment $[c, b_j]$, and b_j is in the boundary of $W(B)$ for each j. We note that there are at least two nonzero y_j's; this is because if otherwise, then we obtain the inequality $k_1(B) \geqslant k_1(B) + 1$, which is a contradiction. Hence we may assume that $y_1, \ldots, y_h \neq 0$, where $h \geqslant 2$, and that this h is the maximum such number.

If c is not in $W(B)$, then there are exactly two points p and q in the boundary of $W(B)$ such that the two line segments $[c, p]$ and $[c, q]$ are in the boundary of $W(A)$ and the relative interior of these two line segments are disjoint from the boundary of $W(B)$ by the fact that $W(A)$ is the convex hull of the union of $W(B)$ and the singleton $\{c\}$. Hence there are three cases to consider: the intersection of the boundary of $W(B)$ and the supporting line at p (resp., q) containing $[c, p]$ (resp., $[c, q]$) is (1) $\{p\}$ (resp., $\{q\}$), (2) a line segment $[p, p']$ (resp., $\{q\}$) or $\{p\}$ (resp., a line segment $[q, q']$), or (3) a line segment $[p, p']$ (resp., a line segment $[q, q']$) (cf. Fig. 1). We need only prove case (2) since other cases can be done similarly.

Define three (disjoint) subsets consisting of the corresponding unit vectors, and their cardinal numbers, respectively, in the following:

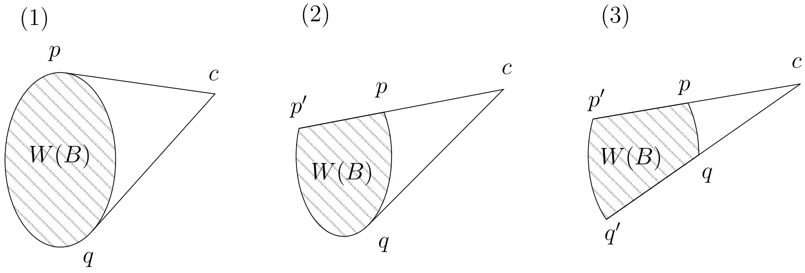

Fig. 1.

$$R \equiv \{z_j\colon \ \xi_j \in [c, p')\} \quad \text{with } r \equiv \#(R),$$
$$S \equiv \{z_j\colon \ \xi_j \in (c, q)\} \quad \text{with } s \equiv \#(S), \quad \text{and}$$
$$T \equiv \{z_j\colon \ \xi_j \in \partial W(A)\backslash([c, p') \cup (c, q))\} \quad \text{with } t \equiv \#(T).$$

So, $k(A) = r + s + t$. Obviously, every $z_j \in T$ is of the form $x_j \oplus 0$. Moreover, we partition R into two disjoint subsets $R_1 \equiv \{z_j\colon \ y_j \neq 0\}$ and $R_2 \equiv \{z_j\colon \ y_j = 0\}$. We call their cardinal numbers r_1 and r_2, respectively. Without loss of generality, we may assume that $R_1 = \{z_1, \dots, z_{r_1}\}$, $R_2 = \{z_{r_1+1}, \dots, z_{r_1+r_2}\}$, $S = \{z_{r+1}, \dots, z_{r+s}\}$, and $T = \{z_{r+s+1}, \dots, z_{r+s+t}\}$, where $r_1 + r_2 = r$. This shows that $r_1 + s = h \geqslant 2$.

First assume that $s = 0$. Then $r_1 \geqslant 2$. For clarity of the proof, the following method is called Method I. Since every y_j, $j = 1, \dots, r_1$, is nonzero, we define the vectors $z'_j = (x_j/y_j) \oplus 1$ for these j's so that the vectors in $M \equiv \{(z'_1 - z'_j)/\|z'_1 - z'_j\|\}_{j=2}^{r_1} = \{(((x_1/y_1) - (x_j/y_j)) \oplus 0)/\|z'_1 - z'_j\|\}_{j=2}^{r_1}$ are linearly independent and are perpendicular to vectors in $T \cup R_2$. This together with [2, Theorem 1] shows that $\operatorname{span}(M) \subseteq \bigcup_{\eta \in [c,p']} K_\eta(A)$ and thus every unit vector in $\operatorname{span}(M)$ is a unit vector corresponding to some $\eta \in \partial W(B)$. Choosing an orthonormal basis $\{v_j \oplus 0\}_{j=2}^{r_1}$ for the subspace $\operatorname{span}(M)$, we deduce from the orthonormality of the vectors in $T \cup R_2 \cup \{v_j \oplus 0\}_{j=2}^{r_1}$ that

$$k_1(B) \geqslant t + r_2 + (r_1 - 1) = r + s + t - 1 = k(A) - 1 \geqslant k_1(B) + 1,$$

which is impossible. Hence we must have $s \geqslant 1$.

If $s = 1$, then $r_1 \geqslant 1$. A similar argument as above yields that

$$k_1(B) \geqslant \begin{cases} t + r_2 + 1 & \text{if } r_1 = 1, \text{ and} \\ t + r_2 + (r_1 - 1) + 1 & \text{if } r_1 \geqslant 2 \end{cases}$$

by considering the orthonormal subsets $T \cup R_2 \cup \{(x_{r+1}/\|x_{r+1}\|) \oplus 0\}$ and $T \cup R_2 \cup \{v_j \oplus 0\}_{j=2}^{r_1} \cup \{(x_{r+1}/\|x_{r+1}\|) \oplus 0\}$, where $\{v_j \oplus 0\}_{j=2}^{r_1}$ is an orthonormal subset of $\operatorname{span}(R_1)$, via Method I on R_1. The above inequalities imply that

$$k_1(B) \geqslant \begin{cases} r + s + t - 1 \geqslant k(A) - 1 \geqslant k_1(B) + 1 & \text{if } r_1 = 1, \text{ and} \\ r + s + t - 1 \geqslant k(A) - 1 \geqslant k_1(B) + 1 & \text{if } r_1 \geqslant 2. \end{cases}$$

This is a contradiction. Hence $s \geqslant 2$.

If $r_1 = 0$, then applying Method I on S, we reach a contradiction since

$$k_1(B) \geqslant t + r_2 + (s - 1) = r + s + t - 1 = k(A) - 1 \geqslant k_1(B) + 1.$$

If $r_1 = 1$, then we obviously have the linear independence of the subset $N \equiv \{(z'_1 - z'_j)/\|z'_1 - z'_j\|\}_{j=r+2}^{r+s} = \{(((x_1/y_1) - (x_j/y_j)) \oplus 0)/\|z'_1 - z'_j\|\}_{j=r+2}^{r+s}$ by applying Method I on S again. Let $\{v_j \oplus 0\}_{j=r+2}^{r+s}$ be an orthonormal basis for the subspace $\operatorname{span}(N)$. Hence

$$k_1(B) \geqslant t + r_2 + (s - 1) + 1 = r + s + t - 1 = k(A) - 1 \geqslant k_1(B) + 1$$

by the orthonormality of the vectors in $T \cup R_2 \cup \{v_j \oplus 0\}_{j=r+2}^{r+s} \cup \{(x_1/\|x_1\|) \oplus 0\}$. This is again a contradiction. If $r_1 \geqslant 2$, then applying Method I on S and R_1, we have the linear independence of the subsets $P \equiv \{(z_1' - z_j')/\|z_1' - z_j'\|\}_{j=r+2}^{r+s} = \{(((x_1/y_1) - (x_j/y_j)) \oplus 0)/\|z_1' - z_j'\|\}_{j=r+2}^{r+s}$ and $Q \equiv \{(z_1' - z_j')/\|z_1' - z_j'\|\}_{j=2}^{r_1} = \{(((x_1/y_1) - (x_j/y_j)) \oplus 0)/\|z_1' - z_j'\|\}_{j=2}^{r_1}$, respectively. Let $\{v_j \oplus 0\}_{j=r+2}^{r+s}$ be an orthonormal basis for $\operatorname{span}(P)$. Then $\operatorname{span}(P) \oplus \operatorname{span}(x \oplus y) = \operatorname{span}(S)$ for some unit vector $x \oplus y$ orthogonal to $\operatorname{span}(P)$. Clearly, x is a nonzero vector; this is because if otherwise, then $0 \oplus y (\in \operatorname{span}(S))$ is orthogonal to $z_1 = x_1 \oplus y_1 (\in R_1)$, which contradicts the fact that y and y_1 are nonzero scalars. Let $\{v_j \oplus 0\}_{j=2}^{r_1}$ be an orthonormal basis for the subspace $\operatorname{span}(Q)$. Then we conclude that the subset $T \cup R_2 \cup \{v_j \oplus 0\}_{j=2}^{r_1} \cup \{v_j \oplus 0\}_{j=r+2}^{r+s} \cup \{(x/\|x\|) \oplus 0\}$ is orthonormal so that

$$k_1(B) \geqslant t + r_2 + (r_1 - 1) + (s - 1) + 1 = r + s + t - 1 = k(A) - 1 \geqslant k_1(B) + 1,$$

which is a contradiction. This completes the proof of case (2).

In case (1), we define three subsets consisting of the corresponding unit vectors, and their cardinal numbers, respectively, as follows:

$$R \equiv \{z_j \colon \xi_j \in [c, p)\} \quad \text{with } r \equiv \#(R),$$
$$S \equiv \{z_j \colon \xi_j \in (c, q)\} \quad \text{with } s \equiv \#(S), \quad \text{and}$$
$$T \equiv \{z_j \colon \xi_j \in \partial W(A) \backslash ([c, p) \cup (c, q))\} \quad \text{with } t \equiv \#(T).$$

As for case (3), we have

$$R \equiv \{z_j \colon \xi_j \in [c, p')\} \quad \text{with } r \equiv \#(R),$$
$$S \equiv \{z_j \colon \xi_j \in (c, q')\} \quad \text{with } s \equiv \#(S), \quad \text{and}$$
$$T \equiv \{z_j \colon \xi_j \in \partial W(A) \backslash ([c, p') \cup (c, q'))\} \quad \text{with } t \equiv \#(T).$$

As before, we partition R (resp., S) into two disjoint subsets $R_1 \equiv \{z_j \colon y_j \neq 0\}$ and $R_2 \equiv \{z_j \colon y_j = 0\}$ (resp., $S_1 \equiv \{z_j \colon y_j \neq 0\}$ and $S_2 \equiv \{z_j \colon y_j = 0\}$). Based on the arguments for case (2), we get a series of contradictions for each individual case. In a similar fashion, we remark that if $A = B \oplus cI_m$, where $c \notin W(B)$, then $k(A) = k_1(B) + k_1(cI_m) = k_1(B) + m$. This remark will be used in the remaining part of the proof.

To complete the proof, we let c be in the boundary of $W(B)$. Assume that $\partial W(B)$ contains no line segment. We infer that $c = b_j = \xi_j$ for $j = 1, \ldots, h$ since these ξ_j's are in the (possibly degenerate) line segment $[c, b_j]$ contained in the boundary of $W(B)$. Define a new vector $z_j' = (x_j/y_j) \oplus 1$ for each $j = 1, \ldots, h$. Then the subset $S \equiv \{(z_1' - z_j')/\|z_1' - z_j'\|\}_{j=2}^{h} = \{(((x_1/y_1) - (x_j/y_j)) \oplus 0)/\|z_1' - z_j'\|\}_{j=2}^{h}$ is linearly independent. Since c is an extreme point of $W(A)$, we have $H_c(A) = K_c(A)$ by [2, Theorem 1] and $\operatorname{span}(S)$ is a subspace of $H_c(A)$. Let $\{v_j \oplus 0\}_{j=2}^{h}$ be an orthonormal basis for $\operatorname{span}(S)$. Then $c = \langle A(v_j \oplus 0), v_j \oplus 0\rangle = \langle Bv_j, v_j\rangle$ is in $\partial W(B)$ for $j = 2, \ldots, h$. Hence

$$k(B) \geqslant (h - 1) + (k(A) - h) = k(A) - 1 \geqslant k(B) + 1.$$

This is a contradiction. So, we may assume that $\partial W(B)$ contains a line segment l such that c belongs to l. If c is not an extreme point of l, then we infer that $c = b_j = \xi_j$ or $\xi_j \in (c, b_j)$ for $j = 1, \ldots, h$ since x_j and y_j are nonzero vectors for these j's. Hence $z_j \in H_c(A)$ for $j = 1, \ldots, h$ by [2, Theorem 1]. Similar arguments show that $H_c(A)$ has an orthonormal subset $\{w_j \oplus 0\}_{j=2}^{h}$. Since $H_c(A) = \bigcup_{\eta \in l} K_\eta(A)$ by [2, Theorem 1], this implies that $w_j \oplus 0 \in K_{\eta_j}(A)$, where $\eta_j \in l$, for $j = 2, \ldots, h$. From $\eta_j = \langle A(w_j \oplus 0), w_j \oplus 0\rangle = \langle Bw_j, w_j\rangle \in l \subseteq \partial W(B)$, where $j = 2, \ldots, h$, we reach a contradiction since

$$k(B) \geqslant (h - 1) + (k(A) - h) = k(A) - 1 \geqslant k(B) + 1.$$

For the remaining part of the proof, let c be an extreme point of l, where l is a line segment on the boundary of $W(B)$. We consider two cases: either (a) there is only one line segment in $\partial W(B)$

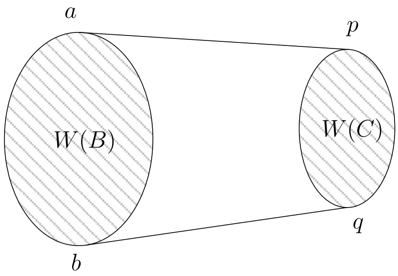

Fig. 2.

passing through c, or (b) there are exactly two line segments in $\partial W(B)$ passing through c. In case (a), since x_j and y_j are nonzero vectors for $j = 1, \ldots, h$, we infer that $c = b_j = \xi_j$ or $\xi_j \in (c, b_j)$ for these j's. This implies that $z_j \in H_\eta(A)$ by [2, Theorem 1], where η is not an extreme point of l. So, the same arguments as above lead us to get a contradiction. For case (b), since c is a corner of $W(B)$, c is a reducing eigenvalue of B by [1, Theorem 1]. Thus B is unitarily similar to a matrix of the form $B' \oplus cI_{n'}$, where c is not an eigenvalue of B', and the size of B' and n' are both less than n. Obviously, $c \notin W(B')$. We apply the preceding remark as for the case of $c \notin W(B)$ to see that $k(A) = k(B' \oplus cI_{n'+1}) = k_1(B') + n' + 1$, and $k(B) = k(B' \oplus cI_{n'}) = k_1(B') + n'$. In addition, $k(B) = k_1(B)$ in this case. Hence we obtain that $k(A) = k_1(B) + 1$, which contradicts our assumption that $k(A) \geqslant k_1(B) + 2$. With this, we conclude the proof of the asserted equality. □

We remark that the part of the proof of Lemma 2.3 on $c \notin W(B)$ involves the following three cases (1), (2), and (3) depending on whether $\partial W(B)$ contains a line segment or otherwise. In case (1), we have $R = \{z_j\colon\ y_j \neq 0\}$ and $S = \{z_j\colon\ y_j \neq 0\}$, in (2) $R = R_1 \cup R_2$, where $R_1 = \{z_j\colon\ y_j \neq 0\}$ and $R_2 = \{z_j\colon\ y_j = 0\}$, and $S = \{z_j\colon\ y_j \neq 0\}$, and in (3) $R = R_1 \cup R_2$, where $R_1 = \{z_j\colon\ y_j \neq 0\}$ and $R_2 = \{z_j\colon\ y_j = 0\}$, and $S = S_1 \cup S_2$, where $S_1 = \{z_j\colon\ y_j \neq 0\}$ and $S_2 = \{z_j\colon\ y_j = 0\}$. Note that the key point is to handle R and S in (1), R_1 and S in (2), and R_1 and S_1 in (3), that is, all nonzero y_j's of the three cases. We find that the proofs of the three cases are almost the same. This observation can facilitate the proof of Theorem 2.2 as follows. If $\partial W(B)$ contains a line segment such that this line segment is a portion of $\partial W(A)$ and stretches to a point of $\partial W(C)$, then we take the same method as the proof of Lemma 2.3 on $c \notin W(B)$ to partition the corresponding R into $R_1 = \{z_j\colon\ y_j \neq 0\}$ and $R_2 = \{z_j\colon\ y_j = 0\}$. As mentioned above, we need only handle R_1. On the other hand, if $\partial W(B)$ contains no such line segments, then we need only handle the corresponding $R = \{z_j\colon\ y_j \neq 0\}$. From this, there is no difference between the proofs of the two cases. Hence we may assume, in the proof of Theorem 2.2, that $\partial W(B)$ and $\partial W(C)$ contain no line segments.

Before giving a proof of Theorem 2.2, we note several things. First of all, by Lemma 2.3, we may assume that both of the numerical ranges $W(B)$ and $W(C)$ are not singletons. Secondly, we may further assume that $\partial W(B)$ and $\partial W(C)$ contain no line segment by the above remark. Thirdly, since $W(A)$ is the convex hull of the union of $W(B)$ and $W(C)$, there are two line segments, called $[a, p]$ and $[b, q]$, in $\partial W(A)$, where $a, b \in \partial W(B)$ and $p, q \in \partial W(C)$. Fourthly, it is easy to check that $a \neq b$ and $p \neq q$. Indeed, if $a = b$, then a is a corner. By [1, Theorem 1], we obtain that a is a reducing eigenvalue of A, and hence a is a reducing eigenvalue of B. This shows that $W(B)$ must contain a line segment, which contradicts our previous assumption. Similarly, we also have $p \neq q$. Combining the above, we have the following Fig. 2 as the numerical range $W(A)$.

As before, by the definition of $k(A)$, there exist $\xi_j = \langle Az_j, z_j\rangle \in \partial W(A)$, $j = 1, 2, \ldots, k(A)$, where $z_j = x_j \oplus y_j$, and $\langle z_i, z_j\rangle = \delta_{ij}$ for $i, j = 1, \ldots, k(A)$. We define four (disjoint) subsets consisting of the corresponding unit vectors, and their cardinal numbers, respectively, as follows:

$$R \equiv \{z_j\colon\ \xi_j \in (a, p)\} \quad \text{with } r \equiv \#(R),$$

$$S \equiv \{z_j\colon\ \xi_j \in (b, q)\} \quad \text{with } s \equiv \#(S),$$

$$T_B \equiv \{z_j\colon\ \xi_j \in \partial W(A) \cap \partial W(B)\} \quad \text{with } t_1 \equiv \#(T_B), \quad \text{and}$$

$$T_C \equiv \{z_j\colon \xi_j \in \partial W(A) \cap \partial W(C)\} \quad \text{with } t_2 \equiv \#(T_C).$$

Since the intersection of $W(B)$ and $W(C)$ is empty, and $\partial W(B)$ and $\partial W(C)$ contain no line segment, we may assume that

$$R = \{z_j = x_j \oplus y_j\colon x_j \neq 0 \text{ and } y_j \neq 0\}_{j=1}^{r},$$
$$S = \{z_j = x_j \oplus y_j\colon x_j \neq 0 \text{ and } y_j \neq 0\}_{j=r+1}^{r+s},$$
$$T_B = \{z_j = x_j \oplus 0\colon x_j \neq 0\}_{j=r+s+1}^{r+s+t_1}, \quad \text{and}$$
$$T_C = \{z_j = 0 \oplus y_j\colon y_j \neq 0\}_{j=r+s+t_1+1}^{r+s+t_1+t_2}.$$

So, $k(A) = r + s + t_1 + t_2$, $k_1(B) \geqslant t_1$ and $k_1(C) \geqslant t_2$. Clearly, the inequality $k(A) \geqslant k_1(B) + k_1(C)$ holds. Now we are ready to prove Theorem 2.2.

Proof of Theorem 2.2. We need only prove that the reversed inequality $k_1(B) + k_1(C) \geqslant k(A)$ holds. First, we consider the case $r = 0$. Assume that $s = 0$. Then our assertion is obvious since

$$k_1(B) + k_1(C) \geqslant t_1 + t_2 = r + s + t_1 + t_2 = k(A).$$

Assume that $s = 1$, i.e., $z_1 = x_1 \oplus y_1 \in S$. Then $k_1(B) \geqslant t_1 + 1$ since the unit vector $(x_1/\|x_1\|) \oplus 0$ is clearly orthogonal to T_B and $\langle B(x_1/\|x_1\|), x_1/\|x_1\|\rangle$ is in $\partial W(B)$ by the convex combination

$$\langle Az_1, z_1\rangle = \|x_1\|^2 \left\langle B\frac{x_1}{\|x_1\|}, \frac{x_1}{\|x_1\|}\right\rangle + \|y_1\|^2 \left\langle C\frac{y_1}{\|y_1\|}, \frac{y_1}{\|y_1\|}\right\rangle \in (b, q).$$

Hence

$$k_1(B) + k_1(C) \geqslant (t_1 + 1) + t_2 = r + s + t_1 + t_2 = k(A).$$

Assume that $s = 2$, i.e., $z_1 = x_1 \oplus y_1$ and $z_2 = x_2 \oplus y_2 \in S$. If x_1 and x_2 are linearly independent, then by the Gram–Schmidt process, there are two unit vectors z_1' and z_2', where $z_j' = x_j' \oplus y_j'$ with $x_j' \neq 0$ for $j = 1, 2$, such that x_1' and x_2' are mutually orthogonal, and $\mathrm{span}(\{z_1, z_2\})$ is equal to $\mathrm{span}(\{z_1', z_2'\})$. Choosing the two unit vectors $(x_1'/\|x_1'\|) \oplus 0$ and $(x_2'/\|x_2'\|) \oplus 0$, we obtain that $k_1(B) \geqslant t_1 + 2$. Hence

$$k_1(B) + k_1(C) \geqslant (t_1 + 2) + t_2 = r + s + t_1 + t_2 = k(A).$$

On the other hand, if x_1 and x_2 are linearly dependent, say, $x_2 = \lambda x_1$ for some scalar λ, then we define a new unit vector

$$z_2' = \frac{z_2 - \lambda z_1}{\|z_2 - \lambda z_1\|} = 0 \oplus \frac{y_2 - \lambda y_1}{\|y_2 - \lambda y_1\|} \in \mathrm{span}\big(\{z_1, z_2\}\big)$$

so that $\mathrm{span}(\{z_1, z_2\}) = \mathrm{span}(\{z_1'\}) \oplus \mathrm{span}(\{z_2'\})$ for some unit vector $z_1' \equiv x_1' \oplus y_1'$, where z_1' and z_2' are mutually orthogonal. Clearly, $x_1' \neq 0$ for otherwise, it leads to $x_1 = x_2 = 0$, which contradicts the definition of S. From the two unit vectors $(x_1'/\|x_1'\|) \oplus 0$ and z_2', we infer that $k_1(B) \geqslant t_1 + 1$ and $k_1(C) \geqslant t_2 + 1$. Hence

$$k_1(B) + k_1(C) \geqslant (t_1 + 1) + (t_2 + 1) = r + s + t_1 + t_2 = k(A).$$

Assume that $s \geqslant 3$, i.e., $S = \{z_j = x_j \oplus y_j\colon x_j \neq 0 \text{ and } y_j \neq 0\}_{j=1}^{s}$. We consider the largest linearly independent subset of $\{x_j\}_{j=1}^{s}$ as follows. Without loss of generality, we may assume that this can be $\{x_j\}_{j=1}^{s}$, $\{x_1\}$ or $\{x_j\}_{j=1}^{l}$, where $1 < l < s$. For the first two cases, it can be done by applying similar arguments as for the case of $s = 2$. In the last case, since x_j is a linear combination of $x_1, \ldots, x_l$ for $j = l + 1, \ldots, s$, it is easy to check that the unit vectors

$$z_j' \equiv \frac{z_j - \sum_{i=1}^{l} a_i^{(j)} z_i}{\|z_j - \sum_{i=1}^{l} a_i^{(j)} z_i\|} = 0 \oplus \left(\frac{y_j - \sum_{i=1}^{l} a_i^{(j)} y_i}{\|y_j - \sum_{i=1}^{l} a_i^{(j)} y_i\|}\right), \quad j = l + 1, \ldots, s, \tag{$*$}$$

are linearly independent. Let $y'_j = \frac{y_j - \sum_{i=1}^{l} a_i^{(j)} y_i}{\| y_j - \sum_{i=1}^{l} a_i^{(j)} y_i \|}$ for $j = l+1, \dots, s$. Since $F \equiv \operatorname{span}(\{z'_j = 0 \oplus y'_j\}_{j=l+1}^{s})$ is a subspace of the space $V \equiv \operatorname{span}(\{z_j\}_{j=1}^{s})$, the orthogonal complement of F in V, called E, can be written as $\operatorname{span}(\{z'_j \equiv x'_j \oplus y'_j\}_{j=1}^{l})$ for some unit vectors z'_j, $j = 1, \dots, l$. By $(*)$, we see that $\{x'_j\}_{j=1}^{l}$ is linearly independent since $\{x_j\}_{j=1}^{l}$ is linearly independent. Hence we may assume that both $\{x'_j\}_{j=1}^{l}$ and $\{y'_j\}_{j=l+1}^{s}$ are orthogonal subsets by the Gram–Schmidt process. This shows that $G_1 \equiv \{(x'_j/\|x'_j\|) \oplus 0\}_{j=1}^{l}$ and $G_2 \equiv \{0 \oplus y'_j\}_{j=l+1}^{s}$ are orthogonal to T_B and T_C, respectively. Since every vector v in G_1 (resp., G_2) is such that $\langle Av, v\rangle$ is in $\partial W(B)$ (resp., $\partial W(C)$), we obtain that $k_1(B) + k_1(C) \geqslant k(A)$ from $k_1(B) \geqslant t_1 + l$ and $k_1(C) \geqslant t_2 + s - l$. This completes the proof of the case $r = 0$.

Next, we prove the case $r = 1$. Obviously, it is sufficient to consider $s \geqslant 1$ since the case $r = 1, s = 0$ is the same as the case $r = 0, s = 1$. Assume that $s = 1$, i.e., $z_1 = x_1 \oplus x_2 \in R$ and $z_2 = x_2 \oplus y_2 \in S$. Then $k_1(B) \geqslant t_1 + 1$ and $k_1(C) \geqslant t_2 + 1$ since $(x_1/\|x_1\|) \oplus 0$ and $0 \oplus (y_2/\|y_2\|)$ are orthogonal to T_B and T_C, respectively. Moreover, $\langle B(x_1/\|x_1\|), x_1/\|x_1\|\rangle$ is in the boundary of $W(B)$ by the convex combination

$$\langle Az_1, z_1\rangle = \|x_1\|^2 \left\langle B\frac{x_1}{\|x_1\|}, \frac{x_1}{\|x_1\|}\right\rangle + \|y_1\|^2 \left\langle C\frac{y_1}{\|y_1\|}, \frac{y_1}{\|y_1\|}\right\rangle \in (a, p),$$

and $\langle C(y_2/\|y_2\|), y_2/\|y_2\|\rangle$ is in the boundary of $W(C)$ by the same arguments. Hence

$$k_1(B) + k_1(C) \geqslant (t_1 + 1) + (t_2 + 1) = r + s + t_1 + t_2 = k(A).$$

Assume that $s = 2$. That is, we have $R = \{z_1 = x_1 \oplus y_1\colon\ x_1 \neq 0 \text{ and } y_1 \neq 0\}$ and $S = \{z_j = x_j \oplus y_j\colon\ x_j \neq 0 \text{ and } y_j \neq 0\}_{j=2}^{3}$. If $\{x_2, x_3\}$ is linearly independent, then we may assume that it is an orthogonal set by the Gram–Schmidt process. By the convex combination mentioned above, we infer from the three unit vectors $0 \oplus (y_1/\|y_1\|)$, $(x_2/\|x_2\|) \oplus 0$, and $(x_3/\|x_3\|) \oplus 0$ that $k_1(B) \geqslant t_1 + 2$ and $k_1(C) \geqslant t_2 + 1$. Hence

$$k_1(B) + k_1(C) \geqslant (t_1 + 2) + (t_1 + 1) = r + s + t_1 + t_2 = k(A).$$

On the other hand, if $\{x_2, x_3\}$ is linearly dependent, say, $x_2 = \lambda x_3$ for some scalar λ, then we define a new unit vector

$$z'_2 = \frac{z_2 - \lambda z_3}{\|z_2 - \lambda z_3\|} = 0 \oplus \frac{y_2 - \lambda y_3}{\|y_2 - \lambda y_3\|} \in \operatorname{span}\bigl(\{z_2, z_3\}\bigr)$$

so that $\operatorname{span}(\{z_2, z_3\}) = \operatorname{span}(\{z'_2\}) \oplus \operatorname{span}(\{z'_3\})$ for some unit vector $z'_3 \equiv x'_3 \oplus y'_3$, where z'_2 is orthogonal to z'_3. Clearly, $x'_3 \neq 0$ for otherwise, it leads to $x_2 = x_3 = 0$, which contradicts the definition of S. From the three unit vectors $0 \oplus (y_1/\|y_1\|)$, $0 \oplus ((y_2 - \lambda y_3)/\|y_2 - \lambda y_3\|)$, and $(x'_3/\|x'_3\|) \oplus 0$, we infer that $k_1(B) \geqslant t_1 + 1$ and $k_1(C) \geqslant t_2 + 2$. Hence

$$k_1(B) + k_1(C) \geqslant (t_1 + 1) + (t_2 + 2) = r + s + t_1 + t_2 = k(A).$$

Assume that $s \geqslant 3$, that is, $S = \{z_j = x_j \oplus y_j\colon\ x_j \neq 0 \text{ and } y_j \neq 0\}_{j=2}^{s+1}$, and $R = \{z_1 = x_1 \oplus y_1\colon\ x_1 \neq 0 \text{ and } y_1 \neq 0\}$. We consider the largest linearly independent subset of $\{x_j\}_{j=2}^{s+1}$ as follows. Without loss of generality, we may assume that this can be $\{x_j\}_{j=2}^{s+1}$, $\{x_2\}$ or $\{x_j\}_{j=2}^{l}$, where $2 < l < s + 1$. The three largest linearly independent subsets are similar to these under $r = 0$, $s \geqslant 3$. Indeed, we need only add this unit vector $0 \oplus (y_1/\|y_1\|)$ to every sub-case of the case $r = 0$, $s \geqslant 3$. Hence we have proved that the reversed inequality $k_1(B) + k_1(C) \geqslant k(A)$. This completes the proof of the case $r = 1$.

Let $r = 2$. With the help of the preceding discussions, we may assume that $s \geqslant 2$. Assume that $s = 2$, that is, $R = \{z_j = x_j \oplus y_j\colon\ x_j \neq 0 \text{ and } y_j \neq 0\}_{j=1}^{2}$ and $S = \{z_j = x_j \oplus y_j\colon\ x_j \neq 0 \text{ and } y_j \neq 0\}_{j=3}^{4}$. If $\{x_3, x_4\}$ is linearly independent, then we consider two cases as follows. First, we assume that $\{y_1, y_2\}$ is linearly independent. We may further assume that $\{x_3, x_4\}$ and $\{y_1, y_2\}$ are orthogonal subsets by the Gram–Schmidt process. Obviously, the two subsets $H_1 \equiv \{0 \oplus (y_1/\|y_1\|),\ 0 \oplus (y_2/\|y_2\|)\}$ and

$H_2 \equiv \{(x_3/\|x_3\|) \oplus 0,\ (x_4/\|x_4\|) \oplus 0\}$ are orthogonal to T_C and T_B, respectively. Since every vector v in H_1 (resp., H_2) is such that $\langle Av, v\rangle$ is in the boundary of $W(C)$ (resp., $W(B)$), we infer, from $k_1(B) \geqslant t_1+2$ and $k_1(C) \geqslant t_2+2$, that $k_1(B)+k_1(C) \geqslant k(A)$. On the other hand, assume that $\{y_1, y_2\}$ is linearly dependent, say, $y_1 = \lambda y_2$ for some scalar λ. Then we define a new unit vector $z_1' = (z_1 - \lambda z_2)/\|z_1 - \lambda z_2\| = ((x_1 - \lambda x_2)/\|x_1 - \lambda x_2\|) \oplus 0$ so that $\operatorname{span}(\{z_1, z_2\}) = \operatorname{span}(\{z_1'\}) \oplus \operatorname{span}(\{z_2'\})$ for some unit vector $z_2' \equiv x_2' \oplus y_2'$, where z_1' and z_2' are mutually orthogonal. Clearly, $y_2' \neq 0$ for otherwise, it leads to $y_1 = y_2 = 0$, which contradicts the definition of R. Moreover, we may assume that $\{x_3, x_4\}$ is an orthogonal subset by the Gram–Schmidt process. Hence $H_3 \equiv \{((x_1 - \lambda x_2)/\|x_1 - \lambda x_2\|) \oplus 0,$ $(x_3/\|x_3\|) \oplus 0,\ (x_4/\|x_4\|) \oplus 0\}$ and $H_4 \equiv \{0 \oplus (y_2'/\|y_2'\|)\}$ are orthogonal to T_B and T_C, respectively. Since every vector v in H_3 (resp., H_4) is such that $\langle Av, v\rangle$ is in the boundary of $W(B)$ (resp., $W(C)$), we infer, from $k_1(B) \geqslant t_1+3$ and $k_1(C) \geqslant t_2+1$, that $k_1(B)+k_1(C) \geqslant k(A)$. On the other hand, if $\{x_3, x_4\}$ is linearly dependent, then we need only consider the case that $\{y_1, y_2\}$ is linearly dependent. So, we may assume that $y_1 = \lambda y_2$ and $x_3 = \mu x_4$ for some scalars λ and μ. Define two new unit vectors

$$z_1' = \frac{z_1 - \lambda z_2}{\|z_1 - \lambda z_2\|} = \frac{x_1 - \lambda x_2}{\|x_1 - \lambda x_2\|} \oplus 0 \quad \text{and} \quad z_3' = \frac{z_3 - \mu z_4}{\|z_3 - \mu z_4\|} = 0 \oplus \frac{y_3 - \mu y_4}{\|y_3 - \mu y_4\|}.$$

Then $\operatorname{span}(\{z_1, z_2\}) = \operatorname{span}(\{z_1'\}) \oplus \operatorname{span}(\{z_2'\})$ and $\operatorname{span}(\{z_3, z_4\}) = \operatorname{span}(\{z_3'\}) \oplus \operatorname{span}(\{z_4'\})$ for some unit vectors $z_2' = x_2' \oplus y_2'$ and $z_4' = x_4' \oplus y_4'$, where z_2' (resp., z_4') is orthogonal to z_1' (resp., z_3'). Clearly, y_2' and x_4' are nonzero by the same argument as above. Hence $H_5 \equiv \{((x_1 - \lambda x_2)/\|x_1 - \lambda x_2\|) \oplus 0,$ $(x_4'/\|x_4'\|) \oplus 0\}$ and $H_6 \equiv \{0 \oplus (y_2'/\|y_2'\|),\ 0 \oplus ((y_3 - \lambda y_4)/\|y_3 - \lambda y_4\|)\}$ are orthogonal to T_B and T_C, respectively. Since every vector v in H_5 (resp., H_6) is such that $\langle Av, v\rangle$ is in the boundary of $W(B)$ (resp., $W(C)$), we infer, from $k_1(B) \geqslant t_1+2$ and $k_1(C) \geqslant t_2+2$, that $k_1(B)+k_1(C) \geqslant k(A)$. Assume that $s \geqslant 3$, i.e., $R = \{z_j = x_j \oplus y_j\colon\ x_j \neq 0 \text{ and } y_j \neq 0\}_{j=1}^{2}$, and $S = \{z_j = x_j \oplus y_j\colon\ x_j \neq 0 \text{ and } y_j \neq 0\}_{j=3}^{s+2}$. If $\{y_1, y_2\}$ is linearly independent, then we may assume that $\{y_1, y_2\}$ is orthogonal by the Gram–Schmidt process. In this case, we consider the largest linearly independent subset of $\{x_j\}_{j=3}^{s+2}$, which may be assumed to be $\{x_j\}_{j=3}^{s+2}$, $\{x_3\}$ or $\{x_j\}_{j=3}^{l}$ $(3 < l < s+2)$. Each of the three cases can be handled by applying similar arguments as for the cases of $r=0$, $s \geqslant 2$. On the other hand, if $\{y_1, y_2\}$ is linearly dependent, say, $y_1 = \lambda y_2$ for some scalar λ, then we define a new unit vector $z_1' = ((x_1 - \lambda x_2)/\|x_1 - \lambda x_2\|) \oplus 0$ so that $\operatorname{span}(\{z_1, z_2\}) = \operatorname{span}(\{z_1'\}) \oplus \operatorname{span}(\{z_2'\})$ for some unit vector $z_2' = x_2' \oplus y_2'$, where z_1' and z_2' are mutually orthogonal. Clearly, y_2' is nonzero by the same argument as for the case of $r=0$, $s=2$. To complete the proof, it remains to consider the three cases mentioned above. By applying similar arguments again as for the cases of $r=0$, $s \geqslant 2$, we obtain the reversed inequality $k_1(B)+k_1(C) \geqslant k(A)$. This completes the proof of the case $r=2$.

Finally, assume that $r \geqslant 3$. It suffices to consider $s \geqslant 3$ since $s \leqslant 2$ has been proven if we exchange the roles of s and r. Hence $R = \{z_j = x_j \oplus y_j\colon\ x_j \neq 0 \text{ and } y_j \neq 0\}_{j=1}^{r}$, and $S = \{z_j = x_j \oplus y_j\colon\ x_j \neq 0 \text{ and } y_j \neq 0\}_{j=r+1}^{r+s}$. As mentioned previously, there are three cases by considering the largest linearly independent subset of $\{y_j\}_{j=1}^{r}$ (resp., $\{x_j\}_{j=r+1}^{r+s}$). Without loss of generality, we may assume that this can be $\{y_j\}_{j=1}^{r}$, $\{y_1\}$ or $\{y_j\}_{j=1}^{l_1}$, where $1 < l_1 < r$, and $\{x_j\}_{j=r+1}^{r+s}$, $\{x_{r+1}\}$ or $\{x_j\}_{j=r+1}^{r+l_2}$, where $1 < l_2 < s$. There are a total of nine cases to be considered. Since each case is similar to the one under $r=0$, $s \geqslant 1$, it follows that the reversed inequality $k_1(B)+k_1(C) \geqslant k(A)$ holds. This completes the proof of the case $r \geqslant 3$. $\square$

At the end of the section, we give a generalization of Theorem 2.2 under a slightly weaker condition on B and C. Let A be a matrix of the form $B \oplus C$. Since $W(A)$ is the convex hull of the union of $W(B)$ and $W(C)$, we consider two (disjoint) subsets of $\partial W(A)$ as follows: one is $\partial W(A) \setminus (\partial W(B) \cup \partial W(C)) \equiv \Gamma_1$, and the other is $\partial W(A) \cap \partial W(B) \cap \partial W(C) \equiv \Gamma_2$. Geometrically, Γ_1 consists of the line segments contained in $\partial W(A)$ but not in $\partial W(B) \cup \partial W(C)$. For Γ_2, since the common boundaries of the three numerical ranges consist of the line segments and points which are not in the line segments, every point of the latter is regarded as a degenerate line segment. Hence Γ_2 consists of the (possibly degenerate) line segments contained in the common boundaries of the three numerical ranges. If $\Gamma \equiv \Gamma_1 \cup \Gamma_2$ consists of at most two (possibly degenerate) line segments, then we say that $W(A)$ has property Λ. Evidently, the disjointness of $W(B)$ and $W(C)$ implies that property Λ holds since Γ_1 consists of exactly two line segments and Γ_2 is empty.

Applying the similar arguments in the proof of Theorem 2.2, property Λ is enough to establish the equality $k(A) = k_1(B) + k_1(C)$. Hence we have the following theorem.

Theorem 2.4. *Let $A = B \oplus C$, where B and C are n-by-n and m-by-m matrices, respectively. If $W(A)$ has property Λ, then $k(A) = k_1(B) + k_1(C) \leqslant k(B) + k(C)$. In this case, $k(A) = k(B) + k(C)$ if and only if $k_1(B) = k(B)$ and $k_1(C) = k(C)$. In particular, $k(A) = m + n$ if and only if $k_1(B) = k(B) = m$ and $k_1(C) = k(C) = n$.*

3. Applications and discussion

The first application of our results in Section 2 is a generalization of Lemma 2.3. Indeed, we are able to determine the value of $k(A)$ for $A = B \oplus C$ with normal C.

Proposition 3.1. *Let $A = B \oplus C$, where C is an m-by-m normal matrix. Then $k(A) = k_1(B) + k_1(C)$. In this case, $k(A) = k(B) + k(C)$ if and only if $k_1(B) = k(B)$ and $k_1(C) = k(C)$. In particular, if $C = cI_m$ for some scalar c, then $k(A) = k_1(B) + k_1(cI_m)$.*

Proof. Let the normal C be unitarily similar to $\bigoplus_{j=1}^{m}[c_j]$. By [7, Lemma 2.9], we may assume that all the c_j's are lying in $\partial W(A)$. This shows that $k_1(C) = m$ immediately. On the other hand, we also obtain $k(A) = k_1(B) + m$ by Lemma 2.3. Hence the asserted equality $k(A) = k_1(B) + k_1(C)$ has been proven. For the remaining part of the proof, it holds trivially by this equality. □

An easy corollary of Proposition 3.1 is the determination of when $k(A)$ equals the size of A for a matrix $A = B \oplus C$ with normal C.

Corollary 3.2. *Let $A = B \oplus C$, where B is an n-by-n matrix and C is an m-by-m normal matrix. Then $k(A) = n + m$ if and only if $k_1(B) = n$ and $k_1(C) = m$. Assume, moreover, that $\dim H_\eta = 1$ for all $\eta \in \partial W(B)$. Then $k(A) = n + m$ if and only if $k_1(B) = n \leqslant 2$ and $k_1(C) = m$.*

Proof. By Proposition 3.1, it is clear that $k(A)$ equals the size of A if and only if $k_1(B)$ and $k_1(C)$ equal the sizes of B and C, respectively. In this case, the assumption on H_η implies that $k_1(B) = n \leqslant 2$ by [7, Proposition 2.10]. This completes the proof. □

For a matrix A of the form $B \oplus C$, we recall the decomposition $\Gamma = \Gamma_1 \cup \Gamma_2$ at the end of Section 2, where $\Gamma_1 = \partial W(A) \setminus (\partial W(B) \cup \partial W(C))$ and $\Gamma_2 = \partial W(A) \cap \partial W(B) \cap \partial W(C)$. The next proposition gives a lower bound for $k(A)$.

Proposition 3.3. *Let $A = B \oplus C$ be an n-by-n $(n \geqslant 3)$ matrix. Then Γ is empty if and only if the numerical range of one summand is contained in the interior of the numerical range of the other summand. In particular, if Γ is nonempty, then $k(A) \geqslant 3$.*

Proof. If $\Gamma = \Gamma_1 \cup \Gamma_2$ is empty, then both Γ_1 and Γ_2 are empty. Since Γ_1 is empty, $\partial W(A)$ is contained in $\partial W(B) \cup \partial W(C)$. This implies that $W(B) \cap W(C)$ is nonempty and thus $W(B) = W(C)$, $W(B) \subseteq \operatorname{int} W(C)$ or $W(C) \subseteq \operatorname{int} W(B)$. Moreover, $\Gamma_2 = \phi$ implies that $W(B) \neq W(C)$. With this, we conclude that either $W(B) \subseteq \operatorname{int} W(C)$ or $W(C) \subseteq \operatorname{int} W(B)$. The converse is obvious. Hence we have proved the first assertion. Let Γ be nonempty, i.e., either Γ_1 or Γ_2 is nonempty. If Γ_1 is nonempty, then there is a line segment on the boundary of $W(A)$. This shows that $k(A) \geqslant 3$ by [7, Corollary 2.5]. On the other hand, if Γ_2 is nonempty, then there is a (possibly degenerate) line segment on the common boundaries of the three numerical ranges. Using [7, Corollary 2.5] again, we may assume that the line segment is degenerate, say, to $\{\xi\}$. This implies immediately that $\dim_\xi H(A) \geqslant 2$. Thus $k(A) \geqslant 3$ by [7, Proposition 2.4]. □

As an application, when A is reducible, the next corollary gives a necessary and sufficient condition for $k(A) = 2$.

Corollary 3.4. *Let $A = B \oplus C$ be an n-by-n $(n \geqslant 3)$ matrix. Then $k(A) = 2$ if and only if either $k(B) = 2$ and $W(C) \subseteq \operatorname{int} W(B)$, or $k(C) = 2$ and $W(B) \subseteq \operatorname{int} W(C)$.*

Proof. If $k(A) = 2$, then Proposition 3.3 shows that Γ is empty, and thus the numerical range of one summand, say, B is contained in the interior of the numerical range of the other summand C. Hence $k(C) = 2$ by [7, Lemma 2.9]. The converse is obvious by [7, Lemma 2.9] again. □

The following proposition determines exactly when $k(A)$ equals the size of A for an irreducible matrix A.

Proposition 3.5. *Let A be an n-by-n $(n \geqslant 3)$ irreducible matrix. Then $k(A) = n$ if and only if $\partial W(A)$ contains a line segment l and there are n points (not necessarily distinct) in $l \cup (\partial W(A) \cap L)$, where L is the supporting line parallel to l, such that their corresponding unit vectors form an orthonormal basis for $\mathbb{C}^n$.*

Proof. We need only prove the necessity. Assume that A is an n-by-n $(n \geqslant 3)$ irreducible matrix with $k(A) = n$. If $\partial W(A)$ contains no line segment, then $\dim H_\xi = \dim E_{\xi,l} \leqslant \frac{n}{2}$ for all $\xi \in \partial W(A)$ by [7, Proposition 2.2]. If n is odd, say, $n = 2m + 1$, then $\dim H_\xi = \dim E_{\xi,l} \leqslant m$ for all $\xi \in \partial W(A)$. Since $k(A) = n$, it follows from [7, Theorem 2.7] that A is reducible, which is absurd. If n is even, say, $n = 2m$, then $m \geqslant 2$ by our assumption that $n \geqslant 3$. Since $k(A) = n$ and $\partial W(A)$ contains no line segment, A is unitarily similar to a matrix of the form

$$\begin{bmatrix} \xi I_m & e^{i\theta} D \\ -e^{i\theta} D^* & \eta I_m \end{bmatrix}$$

by [7, Theorem 2.7], where $\dim H_\xi = \dim H_\eta = m$. Let $D = USV$ be the singular value decomposition of D, where U and V are unitary and $S = \operatorname{diag}(s_1, \ldots, s_m)$ is a diagonal matrix with $s_j \geqslant 0$, $j = 1, \ldots, m$. Then

$$\begin{bmatrix} U^* & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} \xi I_m & e^{i\theta} D \\ -e^{i\theta} D^* & \eta I_m \end{bmatrix} \begin{bmatrix} U & 0 \\ 0 & V^* \end{bmatrix} = \begin{bmatrix} \xi I_m & e^{i\theta} S \\ -e^{i\theta} S & \eta I_m \end{bmatrix}$$

is unitarily similar to

$$\bigoplus_{j=1}^{m} \begin{bmatrix} \xi & e^{i\theta} s_j \\ -e^{i\theta} s_j & \eta \end{bmatrix}.$$

This contradicts the irreducibility of A. Hence $\partial W(A)$ must contain a line segment. We then apply [7, Theorem 2.7] again to complete the proof. □

An easy corollary of Proposition 3.5 is the following upper bound for $k(A)$. This result was shown in [7, Proposition 2.10]. Here we give a simpler proof.

Corollary 3.6. *If A is an n-by-n $(n \geqslant 3)$ matrix with $\dim H_\xi = 1$ for all $\xi \in \partial W(A)$, then $k(A) \leqslant n - 1$.*

Proof. Assume that $k(A) = n$. It suffices to consider that A is reducible; this is because if otherwise, then Proposition 3.5 shows that $\partial W(A)$ contains a line segment, which contradicts the assumption on H_ξ. Let $A = B \oplus C$. Then our assumption on H_ξ implies that Γ is empty. By Proposition 3.3, we obtain that the numerical range of one summand is contained in the interior of the numerical range of the other summand. It follows from [7, Lemma 2.9] that the value of $k(A)$ equals $k(B)$ or $k(C)$. Thus $k(A) \leqslant n - 1$ as asserted. □

We now combine Proposition 3.1, Corollary 3.2, Corollary 3.4, and Proposition 3.5 to determine the value of $k(A)$ for any 4-by-4 reducible matrix A. Corollary 3.4 shows exactly when the value of $k(A)$ equals two. By Proposition 3.1, Corollary 3.2 and Proposition 3.5, we get a necessary and sufficient condition for the value of $k(A)$ to be equal to four. In other words, the value of $k(A)$ can be determined completely for any 4-by-4 reducible matrix A. To do this, we note that a 4-by-4 reducible matrix A can be written, after a unitary similarity, as (i) $A = B \oplus [c]$, where B is a 3-by-3 irreducible matrix and c is a complex number, (ii) $A = B \oplus [c]$, where B is a 3-by-3 reducible matrix and c is a complex number, or (iii) $A = B \oplus C$, where B and C are 2-by-2 irreducible matrices. Proposition 3.7 below is to deal with case (i).

Recall that for a 3-by-3 irreducible matrix A, $W(A)$ is of one of the following shapes (cf. [5]): an elliptic disc, the convex hull of a heart-shaped region, in which case $\partial W(A)$ contains a line segment, and an oval region.

Proposition 3.7. *Let $A = B \oplus [c]$, where B is a 3-by-3 irreducible matrix and c is a complex number. Then $k(A) = 4$ if and only if $c \notin \operatorname{int} W(B)$ and $\{a_1, a_2, b\} \subseteq \partial W(A)$, where $W(B)$ is the convex hull of a heart-shaped region, in which case $\partial W(B)$ contains a line segment $[a_1, a_2]$ contained in the supporting line L_1 of $W(B)$ at a_1 and a_2, and L_2 is the supporting line of $W(B)$ at b which is parallel to L_1.*

Proof. By Corollary 3.2, we see that $k(A) = 4$ is equivalent to $k_1(B) = 3$ and $k_1([c]) = 1$. Since a necessary and sufficient condition for $k_1([c]) = 1$ is that $c \notin \operatorname{int} W(B)$, it remains to show that $k_1(B) = 3$ if and only if $\{a_1, a_2, b\} \subseteq \partial W(A)$ and $W(B)$ satisfies the asserted properties. If $k_1(B) = 3$, then $k(B) = 3$. Hence it follows from Proposition 3.5 that $\partial W(A)$ contains $\{a_1, a_2, b\}$, and $W(B)$ is as asserted. The converse is trivial. □

For case (ii), let $A = B \oplus [c]$, where B is a 3-by-3 reducible matrix. After a unitary similarity, B can be written as $C \oplus [b]$, where C is a 2-by-2 matrix, so that $k(A) = k_1(C) + k_1([b] \oplus [c])$ by Proposition 3.1. The following proposition gives a necessary and sufficient condition for $k(A)$ to be equal to four.

Proposition 3.8. *Let $A = C \oplus [b] \oplus [c]$, where C is a 2-by-2 matrix, and b and c are complex numbers. Then $k(A) = 4$ if and only if both b and c are in $\partial W(A)$ and $k_1(C) = 2$.*

Proof. By Corollary 3.2, it is obvious that $k(A) = 4$ if and only if $k_1(C) = 2$ and $k_1([b] \oplus [c]) = 2$. Moreover, it is also clear that $k_1([b] \oplus [c]) = 2$ is equivalent to both of b and c being in $\partial W(A)$. Hence the proof is complete. □

To prove for case (iii), let $A = B \oplus C$, where B and C are 2-by-2 irreducible matrices. Since $W(A)$ is the convex hull of the union of the two elliptic discs $W(B)$ and $W(C)$, either $W(B)$ equals $W(C)$, or Γ consists of at most four (possibly degenerate) line segments. With this, we are now ready to give a necessary and sufficient condition for $k(A) = 4$.

Proposition 3.9. *Let $A = B \oplus C$, where B and C are 2-by-2 irreducible matrices. Then $k(A) = 4$ if and only if Γ consists of at least three line segments (including the possibly degenerate cases), or Γ consists of exactly two (possibly degenerate) line segments such that $k_1(B) = k_1(C) = 2$.*

Proof. If Γ consists of more than four (possibly degenerate) line segments, then the two elliptic discs $W(B)$ and $W(C)$ are identical. Hence $k(A) = 4$ by direct computations. If Γ consists of four or three (possibly degenerate) line segments, then the endpoints of the major axes of the two elliptic discs $W(B)$ and $W(C)$ are in $\partial W(A)$. Hence $k(A) = 4$. If Γ consists of exactly two (possibly degenerate) line segments such that $k_1(B) = k_1(C) = 2$, then $k(A) = 4$ by Theorem 2.4. Therefore we have proved the sufficient condition for $k(A) = 4$. Next assume that $k(A) = 4$ and either Γ consists of exactly two (possibly degenerate) line segments such that the equalities $k_1(B) = k_1(C) = 2$ fail, or Γ consists of at most one (possibly degenerate) line segment. Since property Λ holds in each case, we must have

$k_1(B) = k_1(C) = 2$ by Theorem 2.4. This shows that we need only consider the latter. If Γ consists of exactly one (possibly degenerate) line segment, then Γ_1 is empty and Γ_2 is a singleton. Hence we may assume that $W(B)$ is contained in $W(C)$ and the intersection of $W(B)$ and $W(C)$ is Γ. This shows that $k_1(B) = 1$ and $k_1(C) = 2$, which is a contradiction. If Γ is empty, then it follows from Proposition 3.3 that the numerical range of one summand, say, B is contained in the interior of the numerical range of the other summand C. By Corollary 3.4 and [3, Lemma 4.1], we see that $k(A) = k(C) = 2$, which is absurd. This completes the proof. □

As a final application of Theorem 2.4, it is obvious that the convex hull of the union of $W(A)$ and $W(A + aI_n)$ has property Λ for any $a \neq 0$. Hence we obtain the following proposition.

Proposition 3.10. *Let A be an n-by-n matrix and a be a nonzero complex number. Then $k(A \oplus (A + aI_n)) = k_1(A) + k_1(A + aI_n)$. In this case, $k(A \oplus (A + aI_n)) = 2k(A)$ if and only if $k_1(A + aI_n) = k_1(A) = k(A)$.*

We conclude this paper by stating the following open questions concerning this topic. Is it true that the equality $k(A) = k_1(B) + k_1(C)$ holds for a matrix A of the form $B \oplus C$ even if property Λ fails? We note that although property Λ fails, the mentioned formula may still be correct (cf. Proposition 3.1). Another natural example of the failure of property Λ is that both $W(B)$ and $W(C)$ have the same numerical range. Is it true that $k(B \oplus C) = k(B) + k(C)$ in this case? In particular, can we determine the value of $k(A \oplus A)$ (cf. Proposition 3.10)? The following proposition gives a partial answer for $k(A \oplus A)$ if we assume, in addition, that $\dim H_\xi = 1$ for all $\xi \in \partial W(A)$.

Proposition 3.11. *If A is an n-by-n matrix with $\dim H_\xi = 1$ for all $\xi \in \partial W(A)$, then*

$$k\left(\bigoplus_{j=1}^{m} A\right) = m \cdot k(A).$$

Proof. Obviously, the inequality $k(\oplus_{j=1}^{m} A) \geqslant m \cdot k(A)$ holds. To prove the reversed inequality, we consider, for convenience, the case $m = 2$. Let $\xi_1 \in \partial W(A \oplus A)$. Then $\dim H_{\xi_1}(A \oplus A) = 2$ by our assumption on $H_\xi(A)$. Hence the subspace $H_{\xi_1}(A \oplus A)$ is spanned by the two unit vectors $x_1 \oplus 0$ and $0 \oplus x_1$, where $\xi_1 = \langle Ax_1, x_1\rangle$. Let z_1 be a unit vector in $H_{\xi_1}(A \oplus A)$. Then $z_1 = (\alpha_1 x_1 \oplus \alpha_2 x_1)/\sqrt{|\alpha_1|^2 + |\alpha_2|^2}$, where α_1 and α_2 are in $\mathbb{C}$. Similarly for $\xi_2 \in \partial W(A \oplus A)$. That is, the subspace $H_{\xi_2}(A \oplus A)$ is spanned by the two unit vectors $x_2 \oplus 0$ and $0 \oplus x_2$, where $\xi_2 = \langle Ax_2, x_2\rangle$. Moreover, if z_2 is a unit vector in $H_{\xi_2}(A \oplus A)$, then $z_2 = (\beta_1 x_2 \oplus \beta_2 x_2)/\sqrt{|\beta_1|^2 + |\beta_2|^2}$, where β_1 and β_2 are in $\mathbb{C}$. Obviously, the orthogonality of z_1 and z_2 is equivalent to $(\alpha_1\bar{\beta}_1 + \alpha_2\bar{\beta}_2)\langle x_1, x_2\rangle = 0$, i.e.,

$$\left\langle \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}, \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \right\rangle \langle x_1, x_2\rangle = 0.$$

This shows that $k(A \oplus A) \leqslant 2k(A)$ immediately by the definition of $k(A)$.

For general m, a similar argument as above yields that

$$\left\langle \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix}, \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} \right\rangle \langle x_1, x_2\rangle = 0$$

for some scalars $\alpha_1, \dots, \alpha_m$ and $\beta_1, \dots, \beta_m$, where x_1 and x_2 are similarly defined. Since the dimension of $\mathbb{C}^m$ is m, the number of these vectors of the form $[\alpha_1, \dots, \alpha_m]^T$ which are orthogonal to each other is at most m. We infer from this and the above equality that the reversed inequality $k(\bigoplus_{j=1}^{m} A) \leqslant m \cdot k(A)$ holds. Therefore we have the asserted equality. □

At the end of this section, we apply Proposition 3.11 to the quadratic matrices. Recall that an n-by-n quadratic matrix A is unitarily similar to a matrix of the form

$$aI_{n_1} \oplus bI_{n_2} \oplus \begin{bmatrix} aI_{n_3} & D \\ 0 & bI_{n_3} \end{bmatrix},$$

where $n_1, n_2, n_3 \geqslant 0$, $n_1 + n_2 + n_3 = n$, $D > 0$, and $a, b \in \sigma(A)$ (cf. [6, Theorem 2.1]).

Corollary 3.12. *Let A be an n-by-n quadratic matrix of the above form. If $D > 0$, then $k(A) = 2 \cdot \#(\{\lambda \in \sigma(D)\colon \lambda = \|D\|\})$.*

Proof. If $D > 0$, then D is unitarily similar to $\operatorname{diag}(d_1, \ldots, d_{n_3})$, where $d_1 = \cdots = d_p = \|D\| \equiv d > d_{p+1} \geqslant \cdots \geqslant d_{n_3} \geqslant 0$ $(1 \leqslant p \leqslant n_3)$. Hence A is unitarily similar to a matrix of the form $aI_{n_1} \oplus bI_{n_2} \bigoplus_{j=1}^{p} B \bigoplus_{j=p+1}^{n_3} B_j$, where $n_1 + n_2 + 2n_3 = n$,

$$B \equiv \begin{pmatrix} a & d \\ 0 & b \end{pmatrix}, \quad \text{and} \quad B_j \equiv \begin{bmatrix} a & d_j \\ 0 & b \end{bmatrix}, \quad j = p+1, \ldots, n_3.$$

Since the set $\{a, b\}$ and all of the numerical ranges $W(B_j)$, $j = p+1, \ldots, m$, are contained in the interior of $W(B)$, it follows from [7, Lemma 2.9] that $k(A) = k(\bigoplus_{j=1}^{p} B)$. Since $\dim H_\xi(B) = 1$ for all $\xi \in \partial W(B)$, we have $k(A) = p \cdot k(B)$ by Proposition 3.11. Obviously, $k(B) = 2$ by [3, Lemma 4.1]. Thus $k(A) = 2p$ as asserted. □

We remark that in the preceding proof the equality $k(\bigoplus_{j=1}^{p} B) = 2p$ can also be established directly. Indeed, the inequality $k(\bigoplus_{j=1}^{p} B) \geqslant 2p$ holds trivially and we can infer from [3, Lemma 4.1] that $k(\bigoplus_{j=1}^{p} B) = 2p$.

References

[1] W.F. Donoghue Jr., On the numerical range of a bounded operator, Mich. Math. J. 4 (1957) 261–263.
[2] M.R. Embry, The numerical range of an operator, Pac. J. Math. 32 (1970) 647–650.
[3] H.-L. Gau, P.Y. Wu, Numerical ranges and compressions of S_n-matrices, Oper. Matrices 7 (2013) 465–476.
[4] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge Univ. Press, Cambridge, 1991.
[5] D.S. Keeler, L. Rodman, I.M. Spitkovsky, The numerical range of 3 × 3 matrices, Linear Algebra Appl. 252 (1997) 115–139.
[6] S.-H. Tso, P.Y. Wu, Matrical ranges of quadratic operators, Rocky Mt. J. Math. 29 (1999) 1139–1152.
[7] K.-Z. Wang, P.Y. Wu, Diagonals and numerical ranges of weighted shift matrices, Linear Algebra Appl. 438 (2013) 514–532.